ELSEVIER

Discrete Applied Mathematics 82 (1998) 263–269

DISCRETE APPLIED MATHEMATICS

Note

On trees and noncrossing partitions

Martin Klazar [*,1]

Department of Applied Mathematics, Charles University, Malostranské náměstí 25, 118 00 Praha 1, Czech Republic

Received 3 December 1996; received in revised form 31 July 1997; accepted 11 August 1997

Abstract

We give a simple and natural proof of (an extension of) the identity $P(k,l,n)=P_2(k-1,l-1,n-1)$. The number $P(k,l,n)$ counts noncrossing partitions of $\{1,2,\ldots,l\}$ into n parts such that no part contains two numbers x and y, $0<y-x<k$. The lower index 2 indicates partitions with no part of size three or more. We use the identity to give quick proofs of the closed formulae for $P(k,l,n)$ when k is 1, 2, or 3. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Noncrossing partition; Enumeration; Tree; Bijection

1. Introduction

By a *partition* we mean a pair $P=(X,L)$, where $X=\{1,2,\ldots,l\}$ is finite ground set and $L\subset\exp(X)$ is a set of nonempty disjoint sets, union of which is X. The elements of L are called *parts*. We use the symbols $|P|$ and $\|P\|$ to refer to the cardinalities $|X|$ and $|L|$, respectively. The symbol (P) denotes the minimum distance $|y-x|$ of two distinct elements of a part of P. In the case that all parts of P are singletons we put $(P)=\infty$. P is said to be *noncrossing* if there are no four distinct numbers $a<b<c<d$ in X and no two distinct parts A,B in L such that $a,c\in A$ and $b,d\in B$. A partition is *poor* if each part has at most two elements.

Let $\mathscr{P}(k,l,n)$ be the set of all noncrossing partitions P for which $|P|=l$, $\|P\|=n$, and $(P)\geqslant k$. Let $\mathscr{P}_2(k,l,n)\subset\mathscr{P}(k,l,n)$ be the subset consisting of poor partitions. Our first goal is to prove via bijection the following identity.

Theorem 1.1. *For any triple of integers $k\geqslant 2$ and $l,n\geqslant 1$ it is true that*

$$|\mathscr{P}(k,l,n)|=|\mathscr{P}_2(k-1,l-1,n-1)|. \tag{1}$$

* E-mail: klazar@kam.ms.mff.cuni.cz.

[1] Supported by Grants GAUK 194 and GAČR 0194.

PII S0166-218X(97)00118-2

The identity appeared first in Simion and Ullman [13] in a coarser version (with two parameters k and l). In [5] we gave a generating functions proof of the present version. In the next section we prove, constructing a bijection matching partitions of both kinds, a further generalization which explains in a natural and simple way why the identity must hold.

Our second goal is to remind the following formulae and to give bijective proofs for them.

Theorem 1.2. *Let $l, n \geqslant 1$ be integers. Then*

$$|\mathscr{P}(1,l,n)| = \frac{1}{l}\binom{l}{n}\binom{l}{n-1}, \tag{2}$$

$$|\mathscr{P}(2,l,n)| = \frac{1}{l-n+1}\binom{2l-2n}{l-n}\binom{l-1}{2l-2n}, \tag{3}$$

and

$$|\mathscr{P}(3,l,n)| = \frac{1}{n-1}\binom{n-1}{l-n+1}\binom{n-1}{l-n}. \tag{4}$$

Formula (2) was proved already by Kreweras [7]. Formula (3) was derived by means of generating functions by Gardy and Gouyou-Beauchamps [3]. Our bijective proof is new and so is the formula (4). However, we will see that (4) is via Theorem 1.1 equivalent to a result of Schmitt and Waterman [10].

We prove (2)–(4) in the beginning of the next section. Our proofs are bijective, the proofs of (3) and (4) use (1). After that we introduce the sequential form of partitions and we prove a general bijective result that implies (1). Before closing this section we want to mention some ideas hidden behind noncrossing partitions. We know of three different motivations which led to noncrossing partitions (or to an equivalent structure).

(1) *Partitions lattice.* All partitions of $\{1, 2, \ldots, l\}$ with the refinement order form a lattice. In the papers [2, 7, 8, 12, 13], … noncrossing partitions are treated from the point of view of that general theory.

(2) *Davenport–Schinzel sequences.* One of the special cases of these sequences is when we forbid in a finite sequence any subsequence of the type *abab*. Then noncrossing partitions arise, only the language is different. Davenport–Schinzel sequences have geometric motivation and they found many applications in computational geometry. Their definition, more information and more references can be found in [3, 4, 6, 11].

(3) *Nucleic acids.* Noncrossing partitions, namely the partitions in $\mathscr{P}_2(k,l,n)$, are used to describe secondary structure of molecules of nucleic acids (e.g. [10, 14]). The reader will find more information and references in [15].

2. The proofs

By a *tree* we mean here a finite rooted tree in which each set of children of a vertex is linearly ordered. We do not distinguish two trees which are isomorphic via an isomorphism of rooted trees preserving the linear orders. It is well known that the number of different trees on n vertices is

$$C_n = \frac{1}{n}\binom{2n-2}{n-1}, \tag{5}$$

the nth Catalan number. There is a well-known bijection between the set of trees with n vertices and the set of proper bracketings with $n-1$ brackets. Hence (5) counts bracketings as well. For a bijective proof of (5) using bracketings see [9]. Another classical enumerative result concerning trees is the formula for the number $N(a,b)$ of different trees with a vertices and b leaves (vertices with no child). Namely,

$$N(a,b) = N(a,a-b) = \frac{1}{a-1}\binom{a-1}{b}\binom{a-1}{b-1}, \tag{6}$$

the Narayana (or Runyon) number. For a bijective proof see [1].

Now, we proceed to the proof of Theorem 1.2. We will see that (2) and (4) reduce to Narayana numbers and (3) to Catalan numbers.

The proof of formula (2). Although this is well known, we include, for the sake of completeness, the reduction to (6). Let $P=(X,L)\in\mathscr{P}(1,l,n)$ and let $x<y$ be two elements of X. We say that x *covers* y if they lie in different parts of P and there is a number $z, z>y$, that lies in the same part as x. We define a graph $T=(V,E)$, where $V=\{0\}\cup X=\{0,1,\ldots,l\}$ and $\{x,y\}$ is an edge iff (a) x and y lie in the same part and are not separated by another element of that part or (b) y is the minimum element of some part and x is the maximum element covering it or (c) y is the minimum element of some part, $x=0$, and no element covers y. Obviously, T is a graph-theoretical tree. Defining 0 to be the root and ordering any set of children of a vertex by the standard ordering of integers, we change T into a tree in our sense.

Clearly, T has $l+1$ vertices and n leaves (the maximum elements of the parts). It is not difficult to revert the procedure and, given a T with $l+1$ vertices and n leaves, to recover the P we started with. We leave this to the interested reader as an exercise. Thus, we have a bijection between $\mathscr{P}(1,l,n)$ and the set of trees with $l+1$ vertices and n leaves. Formula (2) follows from (6). □

The proof of formula (3). In view of (1) it suffices to find the cardinality $|\mathscr{P}_2(1,l-1,n-1)|$. Any partition $P=(X,L)\in\mathscr{P}_2(1,l-1,n-1)$ has $d=l-n$ doubleton parts and $s=2n-l-1$ singleton parts. The doubleton parts form a proper bracketing with d brackets. Thus, by (5), we have $\binom{2d}{d}/(d+1)$ possibilities for the doubleton parts. Singleton parts can be distributed in the $2d+1$ gaps determined by the doubleton parts in an arbitrary way. We have $\binom{2d+s}{s}$ possibilities for the distribution. Now, (3) follows by taking the product and substituting for d and s. □

The proof of formula (4). Again, by (1), it suffices to find the cardinality $|\mathscr{P}_2(2, l-1, n-1)|$. This was done by Schmitt and Waterman in [10], for the readers convenience we repeat their nice simple bijective argument. Let $P=(X,L)\in\mathscr{P}_2(2,l-1,n-1)$. We define a graph $T=(V,E)$, where $V=L\cup\{*\}$ and $\{A,B\}$ is an edge iff (a) x is the maximum element covering y (see the proof of (2)) for some $x\in A, y\in B$ or (b) $A=*$, $y\in B$ and no element covers y. It is easy to see that T is a tree (with the root $*$ and linear orders of children of a vertex induced by the standard linear order of integers). Obviously, T has n vertices and its leaves are exactly the singleton parts of P. Thus, the number of leaves equals $2n-l-1$. It is straightforward to verify that the mapping $P\to T$ is a bijection between $\mathscr{P}_2(2,l-1,n-1)$ and the set of trees with n vertices and $2n-l-1$ leaves. Thus, (4) follows by (1) and (6). □

The rest of the paper is devoted to the proof of (1). We will use a restricted class of trees.

By $\mathscr{T}$ we denote the set of *binary* trees in which each vertex has at most two children. If T is a binary tree and its root has one child, resp. two children, we denote by T_c, resp. by T_ℓ and T_r, the subtree rooted in the child, resp. the subtrees rooted in the left child and in the right child. By the *left tail* of a binary tree T we mean any sequence of vertices $(v_1,v_2,\ldots,v_j)$ such that v_{i+1} is a child of v_i, v_2 is the left child of v_1, and v_j is a leaf. The number j is the *length* of the left tail.

We find it more convenient for the proof to work with partitions expressed as finite sequences. Any sequence $u=a_1a_2\ldots a_l$ of some symbols *determines* a unique partition $P=(X,L)$, where $X=\{1,2,\ldots,l\}$ and i and j are in the same part iff $a_i=a_j$. Given a partition P, there are many sequences $u=a_1a_2\ldots a_l$ determining P. It is clear that the two following requirements – (a) $\{a_1,a_2,\ldots,a_l\}=\{1,2,\ldots,n\}$ and (b) $1\leqslant i<j\leqslant n$ implies that the first i-occurrence in u precedes the first j-occurrence – make u unique. We call finite sequences $u=a_1a_2\ldots a_l$ having properties (a) and (b) *normal sequences.* The process of replacing a sequence v by a normal sequence that determines the same partition is called *normalization*. Thus, any partition $P=(\{1,2,\ldots,l\},L)$ is determined by exactly one normal sequence u. From now on, we work with u instead of P.

We transfer the terminology and notation from partitions to sequences. The symbols $|u|, \|u\|$, and (u) mean the length of u, the number of distinct symbols in u, and the minimum distance of two distinct occurrences of the same symbol in u. That u is poor means that each symbol appears in u at most twice. A (normal) sequence u is noncrossing if it contains no subsequence of the type *abab* (cf. Davenport–Schinzel sequences).

To illustrate the notion of normal sequences we list the partitions which are involved in the instance of (1) $k=2, l=5$ and n unrestricted.

$$\mathscr{P}(2,5,\cdot)=\{12345,12343,12342,12341,12324,12321,12314,12134,12131\},$$

$$\mathscr{P}_2(1,4,\cdot)=\{1234,1233,1232,1231,1223,1221,1213,1123,1122\}.$$

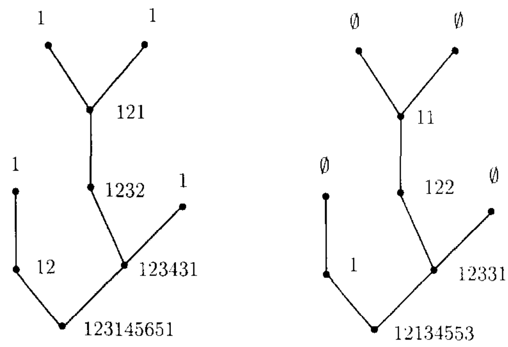

Fig. 1. Mappings F_A and F_B.

We have listed the sequences in the lexicographical order. Looking at them more carefully we notice that not only there is equal number of sequences $u \in \mathscr{P}(2, 5, \cdot)$ with $\|u\| = n$ as sequences $u \in \mathscr{P}_2(1, 4, \cdot)$ with $\|u\| = n - 1$, but that those sequences even appear on the same positions. For instance, for $n = 3$ on the sixth and nineth positions.

And this is the case in general. Before we state the general result we need a few more definitions. We use the standard lexicographical ordering: $u = a_1 a_2 \dots a_m \prec v = b_1 b_2 \dots b_l$ iff either there is an i such that $a_1 = b_1, \dots, a_{i-1} = b_{i-1}, a_i < b_i$ or u is a proper initial segment of v. By $\mathscr{A}$ we denote the set of all normal noncrossing sequences u satisfying $|u| > 0$ (i.e. u is nonempty) and $(u) \geqslant 2$ (i.e. u has no two immediate repetitions of a symbol). By $\mathscr{B}$ we denote the set of all *poor* normal noncrossing sequences u, the empty sequence is included.

The following theorem implies (1) and much more (we mean the phenomenon we have illustrated by the example – sequences with corresponding parameters occupy the same lexicographical positions).

Theorem 2.1. *There is an isomorphism $\Phi : \mathscr{A} \to \mathscr{B}$ of the countably infinite linear orders $(\mathscr{A}, \prec)$ and $(\mathscr{B}, \prec)$ such that, for any $u \in \mathscr{A}$,*

$$|\Phi(u)| + 1 = |u|, \quad \|\Phi(u)\| + 1 = \|u\|, \quad \text{and} \quad (\Phi(u)) + 1 = (u). \tag{7}$$

Proof. We define recursively two mappings $F_A : \mathscr{T} \to \mathscr{A}$ and $F_B : \mathscr{T} \to \mathscr{B}$, see Fig. 1. We use the following notation. If u is a sequence of integers and m an integer then the sequence u^{+m} arises from u by adding m to each term of u. The sequence u^{++m} arises by adding m to each term except for the occurrences of 1.

If T is the one-vertex tree, we set $F_A(T) = 1$ and $F_B(T) = \emptyset$. If the root of T has one child, we set $F_A(T) = 1\ F_A(T_c)^{+1}$ and $F_B(T) = 1\ F_B(T_c)^{-1}$. If the root of T has two children and $a = \|F_A(T_l)\|$ and $b = \|F_B(T_l)\|$, we set $F_A(T) = 1\ F_A(T_l)^{+1}\ F_A(T_r)^{++a}$ and $F_B(T) = 1\ F_B(T_l)^{+1}\ 1\ F_B(T_r)^{+(b-1)}$.

Obviously, F_A maps $\mathscr{T}$ to $\mathscr{A}$ and F_B maps $\mathscr{T}$ to $\mathscr{B}$. To show that F_A and F_B are bijections we invert them. To invert F_A, consider a normal sequence $u \in \mathscr{A}$. If $u = 1$ then $T = F_A^{-1}(u)$ is the one-vertex tree. If $u = 1v$ and 1 does not appear in v, then the root of $T = F_A^{-1}(u)$ has one child and $T_c = F_A^{-1}(w)$ where w is the normalized v. If $u = 1v_1 1v_2$ and 1 does not appear in v_1, then the root of $T = F_A^{-1}(u)$ has two children and $T_\ell = F_A^{-1}(w_1)$ and $T_r = F_A^{-1}(w_2)$, where w_1 is the normalized v_1 and w_2 is the normalized $1v_2$. The mapping F_B can be inverted in a similar way.

Thus $\Phi = F_B \circ F_A^{-1} : \mathscr{A} \to \mathscr{B}$ is a bijection. It is easy to verify by induction on the number of vertices that $|F_A(T)|$ is the number of vertices of T, $\|F_A(T)\|$ is 1+ the number of nonleaf vertices of T, $(F_A(T))$ is the length of the shortest left tail of T, and that these quantities are by one smaller for $F_B(T)$.

Thus, (7) is satisfied. It remains to be proven that Φ is increasing with respect to $\prec$. This will be accomplished by defining a linear ordering $(\mathscr{T}, \prec)$ and showing that both F_A and F_B are increasing mappings.

The linear order on $\mathscr{T}$ is again defined recursively. Let T and U be two distinct binary trees. If T has only one vertex then $T \prec U$. If both T and U have roots with one child, then $T \prec U$ iff $T_c \prec U_c$. If their roots have two children and $T_\ell \neq U_\ell$, then $T \prec U$ iff $T_\ell \prec U_\ell$. If $T_\ell = U_\ell$ then $T \prec U$ iff $T_r \prec U_r$. Finally, let T have root with one child and U with two children. If $T_c \neq U_\ell$ then $T \prec U$ iff $T_c \prec U_\ell$. If $T_c = U_\ell$ then $T \prec U$.

To verify that $(\mathscr{T}, \prec)$ is a linear order and that both mappings are increasing is a matter of a straightforward induction on the number of vertices. As to the linear order, we omit the details. As to the monotonicity, we discuss in detail one case, the other cases are similar. Let $T \prec U$ and let T and U have roots with two children. Let $F_A(T) = t = 1t_1 1t_2$ and $F_A(U) = u = 1u_1 1u_2$, where the splittings of t and u are given by the recursive definition. If $T_\ell \prec U_\ell$ then, by induction, $t_1 \prec u_1$ and this implies $t \prec u$ (note that – for the case when t_1 is a proper initial segment of u_1 – there is no 1 in u_1). If $T_\ell = U_\ell$ and $T_r \prec U_r$ then $1t_1 = 1u_1$ and, by induction, $t_2 \prec u_2$. Again $t \prec u$. □

References

[1] W.Y.C. Chen, A general bijective algorithm for trees, Proc. Nat. Acad. Sci. USA 87 (1990) 9635–9639.
[2] P. Edelman, Chain enumeration and non-crossing partitions, Discrete Math. 31 (1980) 171–180.
[3] D. Gardy, D. Gouyou-Beauchamps, Enumerating Davenport–Schinzel sequences, Inform. théor. et. Appl./Theoret. Inform. Appl. 26 (1992) 387–402.
[4] M. Klazar, Combinatorial aspects of Davenport–Schinzel sequences, Ph.D. Thesis, Department of Applied Mathematics, Charles University, Prague 1995.
[5] M. Klazar, On *abab*-free and *abba*-free set partitions, Eur. J. Combinatorics 17 (1996) 53–68.
[6] M. Klazar, Combinatorial aspects of Davenport–Schinzel sequences, Discrete Math. 165/166 (1997) 431–445.
[7] G. Kreweras, Sur les partitions non croisées d'un cycle, Discrete Math. 1 (1972) 333–350.
[8] Y. Poupard, Etude et denombrement paralleles des partitions non croisées d'un cycle et des decoupages d'un polygone convexe, Discrete Math. 2 (1972) 279–288.
[9] D. Rubenstein, Catalan numbers revisited, J. Combin. Theory A 68 (1994) 486–490.
[10] W.R. Schmitt, M.S. Waterman, Linear trees and RNA secondary structure, Discrete Appl. Math. 51 (1994) 317–323.
[11] M. Sharir, P.K. Agarwal, Davenport–Schinzel Sequences and Their Geometric Applications, Cambridge University Press, Cambridge, 1995.

[12] R. Simion, Combinatorial statistics on non-crossing partitions, J. Combin. Theory A 66 (1994) 270–301.

[13] R. Simion, D. Ullman, On the structure of the lattice of noncrossing partitions, Discrete Math. 98 (1991) 193–206.

[14] P.R. Stein, M.S. Waterman, On some new sequences generalizing the Catalan and Motzkin numbers, Discrete Math. 26 (1979) 261–272.

[15] M.S. Waterman, Applications of combinatorics to molecular biology, in: R.L. Graham, M. Grötschel, L. Lovász (Eds.), Handbook of Combinatorics, North-Holland, Amsterdam, 1995, pp. 1983–2001.